

PROBING THE EXPANSION HISTORY OF THE UNIVERSE BY MODEL-INDEPENDENT RECONSTRUCTION FROM SUPERNOVAE AND GAMMA-RAY BURSTS MEASUREMENTS

CHAO-JUN FENG, XIN-ZHOU LI

Shanghai United Center for Astrophysics (SUCA),
 Shanghai Normal University, 100 Guilin Road, Shanghai 200234, P.R.China
 Email: fengcj@shnu.edu.cn, kychz@shnu.edu.cn

Draft version April 8, 2016

ABSTRACT

To probe the late evolution history of the Universe, we adopt two kinds of optimal basis systems. One of them is constructed by performing the principle component analysis (PCA) and the other is build by taking the multidimensional scaling (MDS) approach. Cosmological observables such as the luminosity distance can be decomposed into these basis systems. These basis are optimized for different kinds of cosmological models that based on different physical assumptions, even for a mixture model of them. Therefore, the so-called feature space that projected from the basis systems is cosmological model independent, and it provide a parameterization for studying and reconstructing the Hubble expansion rate from the supernova luminosity distance and even gamma-ray bursts (GRBs) data with self-calibration. The circular problem when using GRBs as cosmological candles is naturally eliminated in this procedure. By using the Levenberg-Marquardt (LM) technique and the Markov Chain Monte Carlo (MCMC) method, we perform an observational constraint on this kind of parameterization. The data we used include the "joint light-curve analysis" (JLA) data set that consists of 740 Type Ia supernovae (SNIa) as well as 109 long gamma-ray bursts with the well-known Amati relation.

Subject headings: cosmology: cosmological parameters – methods: data analysis

1. INTRODUCTION

One of the major target for present observations is to learn the evolution history of the Universe through the cosmic expansion rate. The observations of the Type Ia supernovae (SNIa) have indicated that the Universe is currently accelerating (Riess et al. 1998; Perlmutter et al. 1999). For the lack of deeper understanding, the cause of this acceleration is usually explained by introducing an exotic energy component called dark energy. There are many dark energy models that based on different physical origins, see Li et al. (2011) for a recent review. A specific dark energy model is usually characterized by a small set of parameters. One can constrain these parameters by observational data to obtain the expansion rate of the Universe. Although this approach is reasonable, the result is often depending on which model one used. So an interesting question is that how to probe the cosmic evolution history from observations without any reference to a specific dark energy model.

All such researches are often called the cosmological model-independent reconstruction of the cosmic expansion rate from observations, and it has been largely discussed in the literature. Most of them are based on a smoothing procedure in redshift bins (Huterer & Starkman 2003). Crittenden & Pogosian (2005); Simpson & Bridle (2006) have also performed PCA to reconstruct the dark energy equation of state. Mignone & Bartelmann (2008) has expanded the luminosity distance into a series of orthonormal functions as basis to reconstruct the cosmic expansion rate. Maturi & Mignone (2009) has optimized this basis system to be capable to describe cosmologies independently of their background physics. The quality of the estimation of the luminosity distance is also improved. Li et al. (2014) has applied this method to determine the curvature parameter.

Gamma-Ray Bursts (GRBs) are the most intense explosions in the Universe, and they can potentially be another standard

candles living in the high redshifts. There are many GRBs observed at $0.1 < z \leq 8.1$, whereas the maximum redshift could to be 10 or even larger in the future observations. So, GRBs is a complementary probe to SNIa, see Schaefer (2007) for a review on the so-called GRB cosmology. However, there is a circularity problem when using GRBs as cosmological candles, because low-redshift GRBs at $z < 0.1$ are too few to calibrate the correction relation in a model-independent way. Then, an input cosmology is needed to obtain the relation, but it leads to the circular problem when constraining cosmological parameters. To alleviate the circularity problem, some statistical methods have been proposed in Ghirlanda et al. (2004), such as the scatter method, the luminosity distance method, and the Bayesian method in Firmani et al. (2005). Liang et al. (2008); Kodama et al. (2008) have suggested calibrating GRBs by using the SNIa data, see also Wei (2010) for a relevant work. Another interesting approach was proposed by Li et al. (2008), in which they have treated the parameters involved in GRBs as free parameters and determined them simultaneously with other cosmological parameters by global fitting.

In this paper, we adopt two kinds of optimal basis systems to probe the evolution history of the Universe. One of them is constructed by performing the principle component analysis (PCA) following the way of Mignone & Bartelmann (2008); Maturi & Mignone (2009). But there are some differences, which will be discussed in the next section. The other kind of optimal basis is build by taking the multidimensional scaling (MDS) approach (Borg et al. 2013, chapter 5), which is another powerful method to reconstruct the cosmic expansion rate. These basis have been optimized for different kinds of cosmological models that based on different physical assumptions, even for a mixture model of them. Therefore, the so-called feature space that projected from the basis systems is cosmological model independent, and it provide a parameterization for studying and reconstructing the Hubble expansion rate from the supernova luminosity distance and even gamma-ray bursts (GRBs) data with self-calibration. The circular problem when using GRBs as cosmological candles is naturally eliminated in this procedure. By using the Levenberg-Marquardt (LM) technique and the Markov Chain Monte Carlo (MCMC) method, we perform an observational constraint on this kind of parameterization. The data we used include the "joint light-curve analysis" (JLA) data set that consists of 740 Type Ia supernovae (SNIa) as well as 109 long gamma-ray bursts with the well-known Amati relation.

terization for studying and reconstructing the Hubble expansion rate from the supernova luminosity distance and even gamma-ray bursts (GRBs) data with self-calibration. By using the Levenberg-Marquardt (LM) technique and the Markov Chain Monte Carlo (MCMC) method, we perform an observational constraint on this kind of parameterization. The data we used include the "joint light-curve analysis" (JLA) data set that consists of 740 Type Ia supernovae (SNIa) as well as 109 long GRBs with the well-known Amati relation. The circular problem when using GRBs as cosmological candles is naturally eliminated in this procedure. This may look like the global fitting method proposed by Li et al. (2008), but here we do not assume any cosmological models in advance.

The structure of this paper is as follows. In Section 2, we present the essential parts of the model-independent method and show how efficient when these methods are applied to optimize the basis for different kinds of cosmological models. The description of data and application of the method to reconstruct the evolution of the Universe are shown in Section 3. The discussions conclusions are presented in Section 4.

2. MODEL-INDEPENDENT METHOD

In this section, we will present applications of PCA and MDS to the JLA and GRB data to probe the evolution of the Universe. At first, we will give the basic formulae and expand cosmological observables into a finite sums of functions as basis. Then, the basis are optimized by using the PCA and MDS methods respectively. In fact, the basic ideas of PCA and MDS are very similar and one can finally obtain the most important components that could be used to describe the observables. It should be noticed that we will directly focus on the cosmological observables like distances instead of the physical quantities within a specific cosmological model, such as the equation of state of dark energy.

2.1. Basic formulae for the cosmic expansion

In Friedmann-Robertson-Walker metric, the luminosity distance is given by

$$D_L(z) = \frac{c}{H_0} \frac{1+z}{\sqrt{|\Omega_k|}} \text{sinn} \left(\sqrt{|\Omega_k|} \int_0^z \frac{dz'}{E(z')} \right), \quad (1)$$

with $E(z) = H(z)/H_0$, and $\text{sinn}(x) = \sin(x), x, \sinh(x)$ for $k = 1, 0, -1$ respectively. Here, c is the speed of light, and $\Omega_k \equiv -k/(a_0 H_0)^2$ denotes the density of the spatial curvature at present. By taking the derivative of Eq.(1) with respect to the redshift z , one can obtain

$$E(z)^{-1} = \frac{D'(z)}{\sqrt{1 + \Omega_k D^2(z)}}, \quad (2)$$

where $D(z)$ is the H_0 -independent comoving angular diameter distance that relates to the luminosity distance as

$$D(z) = \frac{H_0}{c} \frac{D_L(z)}{(1+z)}, \quad (3)$$

and where the prime denotes the derivative with respect to the redshift z . For a flat universe, Eq.(2) could be written as

$$E(z)^{-1} = \frac{H_0}{c} \left[\frac{D'_L(z)}{1+z} - \frac{D_L(z)}{(1+z)^2} \right], \quad (4)$$

Obviously, if the behaviours of both $D(z)$ (or D_L) and its derivative could be dug up from some observational data, one

can obtain the evolution history of the universe from Eq.(2) or (4). Although the data from observations of SNe Ia provide measurements of the distance modulus and redshifts, it is not a convenient way to taking derivative to the luminosity distance directly from the data, because the result would be extremely noisy and unreliable. Therefore, we need to first properly smooth the data by fitting an adequate function $D(z)$ to the measurements in a model-independent way. The derivative can then be approximated by the derivative of $D(z)$. This can be achieved through an expansion of $D(z)$ into a finite sums of suitable functions $p_i(z)$ like:

$$D(z) = \sum_{i=1}^M c_i p_i(z). \quad (5)$$

The M coefficients c_i can be determined by fitting the data, namely c_i are those which minimize the χ^2 statistic function. The number of the terms to be included in the expansion depends on the choice of the orthonormal basis and the quality of the data. The basis $\{p_i\}$ could be arbitrary with idea data, but it will not be in practice. Benitez-Herrera et al. (2012) has used the Gram-Schmidt orthonormalization to decompose the luminosity distance, and they found a systematic trend on the slope of the reconstructed cosmic expansion rate. It indicated that a randomly chosen system of orthonormal basis functions may not be well adapted to the behavior of the measured data. Maturi & Mignone (2009) has suggested optimizing the basis system by using of PCA to reduce the number of coefficients M in Eq.(5), and the possible bias introduced by the choice of the basis is also removed, see Benitez-Herrera et al. (2013). In this paper, we will make use of two optimal basis systems that one derived from the principal component analysis (PCA) and the other one from the multidimensional scaling (MDS) approach. The number of coefficients required is minimized by either PCA or MDS methods. Besides, it also removes any bias introduced by the choice of the basis.

2.2. The Optimal Basis

2.2.1. The Training Set and its Generator

To obtain the optimal basis, we start by writing the H_0 -independent comoving angular diameter distance in a column vector $\mathbf{D} = (D(z_1), D(z_2), \dots, D(z_n))^T \in \mathbb{R}^n$, which can be regarded as a single point in an n -dimensional space. In the literature, n is often taken to be the number of data points from observations, but we will expand the variables at the redshifts in a certain range with a small interval, say 0.1. And then we apply the spline interpolation method to calculate the distance at data points. Also, the range of the redshifts is enlarged to cover that of the GRBs.

Now we select a group of models that are believed to space the set of variable cosmologies and calculate \mathbf{D} for each model to generate a set of vectors \mathbf{D}_i with $i = 1, 2, \dots, M$, where M is the number of models. The ensemble of models $\mathbf{T} = (\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_M) \in \mathbb{R}^{n \times M}$ are called the *training set* introduced by Maturi & Mignone (2009). In principle, the train set could be constructed from any models with arbitrary functions, but it is convenient to consider the models at least weakly resembling the data set (Maturi & Mignone 2009). In other words, one can choose any models, as long as the data set is tightly enclosed in the distribution of the M -point cluster in the n -dimensional space (Maturi & Mignone 2009; Benitez-Herrera et al. 2013; Li et al. 2014). To avoid confusion with a specific cosmological model that determines the

evolution of the universe, we would like to call these models the *training set generators* (TSGs), which mean they are only responsible for building the training set.

In the literature, the Λ CDM model with parameters uniformly sampled in the parameter space are often considered as a TSG to build the training set, but of course other kinds of cosmological models can be used as well, such as the dynamical dark energy models, modified gravity models, or even a mixture of them. However, the result optimal basis system is independent of any TSGs. To see this, we will take the non-flat Λ CDM, the w CDM, the Chevallier-Polarski-Linder (CPL) parametrization model (Chevallier & Polarski 2001; Linder 2003), the FSLP parametrization model without divergence (Feng et al. 2012), the holographic dark energy (HDE) model (Li 2004), the Dvali-Gabadadze-Porrati (DGP) model (Dvali et al. 2000; Deffayet 2001, 2002), the new agegraphic dark energy (NADE) model (Wei & Cai 2008), the Ricci dark energy (RDE) model (Gao et al. 2009; Feng & Li 2009) and their mixture as the TSGs. It also shows that no matter which kind of TSG we used, the dimensionality of the training set could be reduced efficiently by using either the principal component analysis (PCA) or the multidimensional scaling analysis (MDS).

2.2.2. Building the Optimal Basis with PCA

PCA is a very useful statistical tool to reduce the dimensionality of an initially large training set space. Taking the mean of the training set, we obtain a reference model \mathbf{D}_{ref} that defines the origin of the n -dimensional space:

$$\mathbf{D}_{\text{ref}} = \langle \mathbf{D}_i \rangle = \frac{1}{M} \sum_{i=1}^M \mathbf{D}_i \in \mathbb{R}^{n \times 1}. \quad (6)$$

Then, one can define the so-called *covariance matrix* by :

$$\mathbf{S} = \frac{1}{M} \Delta \Delta^T, \quad \text{with} \quad \Delta = \mathbf{T} - \mathbf{D}_{\text{ref}} A \in \mathbb{R}^{n \times M}, \quad (7)$$

where $A = (1, 1, 1, \dots, 1) \in \mathbb{R}^{1 \times M}$. Therefore, the principal components (PCs) are the eigenvectors of the \mathbf{S} matrix, which can be obtained by solving the eigenvalue problem $\mathbf{S} \mathbf{w}_i = \lambda_i \mathbf{w}_i$. In the following, the eigenvalues $\lambda_i (i = 1, 2, \dots, n)$ are sorted in a descendent sequence $\lambda_i > \lambda_{i+1}$, and the corresponding eigenvectors \mathbf{w}_i are called the first PC (\mathbf{w}_1), the second PC (\mathbf{w}_2), and so on. This gives us the components in order of significance, and we can decide to ignore the components of lesser significance. For instance, if we choose only the first p eigenvectors, then the information content of the training set can be optimised via a linear transformation \mathbf{W} : $\mathbb{R}^n \rightarrow \mathbb{R}^p$ mapping the training set vectors into a so called *feature space*: $\mathbf{t}_i = \mathbf{W}^T \mathbf{D}_i \in \mathbb{R}^p, (i = 1, 2, \dots, M)$. Here, \mathbf{t}_i are called the *feature vectors*, while the linear transformation is given by: $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p)$. We do lose some information for ignoring $(\mathbf{w}_{p+1}, \mathbf{w}_{p+2}, \dots, \mathbf{w}_n)$, but if their eigenvalues are small enough, we do not lose much. Then, one could expand $D(z)$ into the optimal basis as

$$D(z) = \mathbf{D}_{\text{ref}} + \sum_{i=1}^p c_i \mathbf{w}_i, \quad (8)$$

with some coefficients c_i that will be determined by fitting data through χ^2 minimization. The $D'(z)$ is derived by taking derivative with respect to the redshift on both side of Eq.(8). Since the eigenvalue λ_i is just the variance of Δ along the

vector \mathbf{w}_i , the percentage of variance we are willing to consider will then determine the number of PCs to be included in the reconstruction matrix \mathbf{W} , i.e. the value of p . For example, we define the cumulative percentage of total variation (Benitez-Herrera et al. 2013; Jolliffe 2002, section 6.1.1) as:

$$r_p = \frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^n \lambda_i}, \quad (9)$$

and after setting a threshold, e.g. $r_p > 99\%$, it will return the value of p .

2.2.3. Building the Optimal Basis with MDS

MDS is another useful statistical tools to reduce the dimensionality of the training set. There are many types of MDS (Borg et al. 2013, chapter 5), which can be classified according to whether the similarities data are qualitative (called non-metric MDS) or quantitative (called metric MDS). In this paper, we will take the algorithms of so-called the classical MDS (CMDS), a special kind of metric MDS. In CMDS, a single Euclidean distance matrix is often used. From the training set \mathbf{T} built before, one can easily construct a square-distance matrix \mathbf{Q} (Borg et al. 2013, chapter 5), whose components are given by

$$Q_{ij} = \sum_{k=1}^n (D_{ik} - D_{jk})^2 \in \mathbb{R}^{M \times M}, \quad (10)$$

with $i = 1, 2, \dots, M$. The matrix \mathbf{Q} describes the dissimilarity of a pair of \mathbf{D} s. Centering the matrix \mathbf{Q} , we obtain the Gram matrix of \mathbf{Q} :

$$\mathbf{G} = -\frac{1}{2} \mathbf{Z} \mathbf{Q} \mathbf{Z}, \quad (11)$$

where $\mathbf{Z} = \mathbf{I}_M - \mathbf{M}^{-1} \mathbf{1} \mathbf{1}^T$ with \mathbf{I}_M the identity matrix of order M , and $\mathbf{1}$ a vector with a 1 in each of its entries. Then, we compute the eigenvalues and eigenvectors of the matrix \mathbf{G} , λ_i, y_i . And as before, λ_i are sorted in a descendent sequences $\lambda_i > \lambda_{i+1}$. Therefore, by taking the first p positive eigenvalues and the corresponding first p eigenvectors, we get the MDS configuration with low dimension $p < M$ as $\mathbf{X} = \mathbf{Y}_+ \Lambda_+^{1/2} \in \mathbb{R}^{M \times p}$, where $\mathbf{Y}_+ = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_p)$ and $\Lambda_+ = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$. Here, $\mathbf{X}^T \in \mathbb{R}^{p \times M}$ plays the same role of the feature space in the PCA mapping, i.e. $\mathbf{t} = \tilde{\mathbf{W}}^T \mathbf{T} \in \mathbb{R}^{p \times M}$. Finally, one could expand $D(z)$ into

$$D(z) = \sum_{i=1}^p \tilde{c}_i \tilde{\mathbf{w}}_i, \quad (12)$$

where the optimal basis are given by

$$\tilde{\mathbf{W}} = (\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \dots, \tilde{\mathbf{w}}_p) = \mathbf{T} \mathbf{T}^T \mathbf{T} \mathbf{Y}_+ \Lambda_+^{1/2} \Sigma \in \mathbb{R}^{n \times p}. \quad (13)$$

Here Σ is a diagonal matrix to rescale the basis, such that the maximum absolute value of each $\tilde{\mathbf{w}}_i$ is equal to one. We shall find that this model is also well consistent with observations. Defining the following two cumulative quantities

$$r_p^{(1)} = \frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^n |\lambda_i|}, \quad r_p^{(2)} = \frac{\sum_{i=1}^p \lambda_i^2}{\sum_{i=1}^n \lambda_i^2}, \quad (14)$$

we can determine the value of p by either of the thresholds is satisfied, e.g. $r_p^{(1)} > 99\%$, or $r_p^{(2)} > 99\%$.

2.3. Efficiency of Different TSGs

For each TSG mentioned before, we construct the training set for 20 times and average the values of r_1 and r_2 for the PCA method, and the values of $r_1^{(1,2)}$ and $r_2^{(1,2)}$ for the MDS method. In each training set, there are 20 models except for the mixture one. In each model, the range of the redshift is $z \in [0, 10]$ with an interval of 0.1, and the parameters of the TSGs are uniformly sampled with boundaries listed as the last column of Tab.2 to calculate the distance \mathbf{D}_i , ($i = 1, 2, \dots, 20$), so that the training set $\mathbf{T} \in \mathbb{R}^{100 \times 20}$. Results are summarized in Tab.2. It should be noticed that the training set from the mixture TSG contains 100 models that randomly chosen from the other TSGs in Tab.2 with the same parameters' ranges. From Tab.2, one can see that the value of p satisfying Eq.(9) or (14) could be very small, say, $p = 1, 2$, whenever which TSG is used.

We have chosen the Λ CDM model as the TSG to build the training set for 10000 times and plotted the distribution of the values of r_1 in the top-left panel of Fig.1 for the PCA analysis, and for each time the number of models is uniformly sampled from 20 to 50. It shows that the first principal component (PC) has the largest possible variance, namely that it retains $> 99.0\%$ of the total variance in the sample. For comparison, the histogram of the second PC in the percentage of total variation, i.e. $r_2 - r_1$ is also plotted in the top-right panel of Fig.1. Therefore, the first two PCs retain $> 99.99\%$ of the total variance, which means that they have already considered the major properties in the expansion of the training set. For the MDS approach, we get almost the same results when using the first kind of threshold in Eq.(14), see the two top panels of Fig.2. However, when the second kind of threshold in Eq.(14) is applied, r_1 has already preserved $> 99.99\%$ of the total squared of the eigenvalues, see Tab.2.

The bottom-left panel of Fig.1 depicts the first four PCs for the comoving angular diameter distances, while the bottom-right one shows the scree plot. It is clear that the feature space with 2-dimension are enough to describe the distance space without losing much information. The same conclusions could be drawn from Fig.2. In fact, by using the MDS approach, the feature space with 1-dimension is good enough.

3. THE COSMIC EXPANSION HISTORY RECONSTRUCTION

Since the feature spaces discussed before retain all significant cosmological information, they can be used to parameterize cosmologies, Maturi & Mignone (2009) called the principal components cosmological eigen-modes (*eigen-cosmologies*). They aim to describe observable quantities directly, while the "standard" cosmological parameters describe the physical properties. To make a distinction between the parameterization from PCA method Eq.(8) and that from MDS method Eq.(12), we will call them the *PCA-model* and the *MDS-model* respectively. In the following, these two models are fitted by the SNeIa and GRBs data. Then, the evolution history of the Universe is obtained by using these two models.

3.1. Data Descriptions

3.1.1. JLA Supernovae Data

The latest large SNeIa data set is the "joint light-curve analysis" (JLA) sample, in which it contains 740 spectroscopically confirmed type Ia supernovae covering the redshift range $0.01 < z < 1.3$ with high quality light curves. The distance

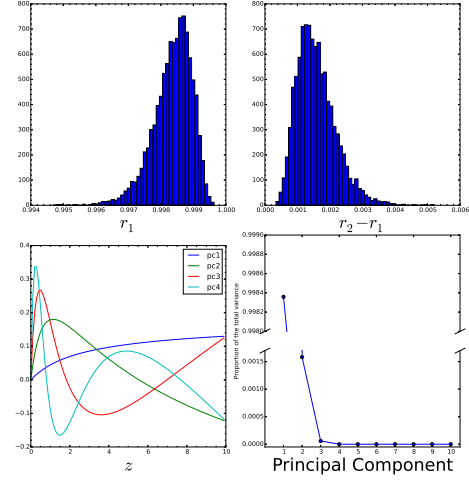


FIG. 1.— PCA method. Top: Histograms of the first PC (left) and second PC (right) in the percentage of total variation from the Λ CDM TSG. Bottom-left: The first 4 PCs for the comoving angular diameter distances. Bottom-right: The scree plot. All the values of parameters are uniformly sampled with boundaries $0.1 < \Omega_m < 0.9, 0.1 < \Omega_\Lambda < 0.9$ and $-0.1 < \Omega_k < 0.1$ as listed in the Tab.2. The range of the redshift is $z \in [0, 10]$ with an interval of 0.1.

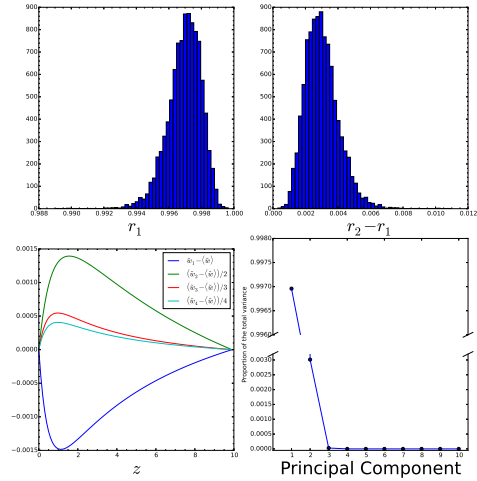


FIG. 2.— MDS method. Top: Histograms of the first eigenvalue (left) and second eigenvalue (right) in the percentage of total absolute eigenvalues from the Λ CDM TSG. Bottom-left: The first 4 basis for the comoving angular diameter distances. To get a better visualization to see how they differ one from the other, we plot their difference with their average and then divided by their number: $(\bar{w}_i - \langle \bar{w} \rangle)/i$ for $i = 1, 2, 3, 4$. Bottom-right: The scree plot. All the values of parameters are uniformly sampled with boundaries $0.1 < \Omega_m < 0.9, 0.1 < \Omega_\Lambda < 0.9$ and $-0.1 < \Omega_k < 0.1$ as listed in the Tab.2. The range of the redshift is $z \in [0, 10]$ with an interval of 0.1.

estimator in this analysis assumes that supernovae with identical color, shape and galactic environment have on average the same intrinsic luminosity for all redshifts. This hypothesis is quantified by a linear model, yielding a standardized distance modulus (Betoule et al. 2014; Shafer 2015)

$$\mu_{\text{obs}} = m_B - (M_B - A \cdot s + B \cdot C + P \cdot \Delta_M), \quad (15)$$

where m_B is the observed peak magnitude in rest-frame B band, M_B, s, C are the absolute magnitude, stretch and color

measures, which are specific to the light-curve fitter employed, and $P(M_* > 10^{10} M_\odot)$ is the probability that the supernova occurred in a high-stellar-mass host galaxy. The stretch, color, and host-mass coefficients (A, B, Δ_M , respectively) are nuisance parameters that should be constrained along with other cosmological parameters. On the other hand, the distance modulus predicted from a cosmological model for a supernova at redshift z is given by

$$\mu_{\text{model}}(z, \vec{\theta}) = 5 \log_{10} \left[\frac{D_L(z)}{10 \text{ pc}} \right], \quad (16)$$

where $\vec{\theta}$ are the cosmological parameters in the model, and $D_L(z)$ is the luminosity distance. For a given pair of the heliocentric-frame and the CMB-frame redshifts ($z_{\text{hel}}, z_{\text{cmb}}$) from the JLA data,

$$\begin{aligned} D_L(z = z_{\text{cmb}}) &= \frac{c}{H_0} \frac{1 + z_{\text{hel}}}{\sqrt{|\Omega_k|}} \text{sinn} \left(\sqrt{|\Omega_k|} \int_0^{z_{\text{cmb}}} \frac{dz'}{E(z')} \right) \\ &= (1 + z_{\text{hel}}) r_A(z_{\text{cmb}}), \end{aligned} \quad (17)$$

where $r_A(z)$ is the comoving angular diameter distance. The χ^2 statistic is then calculated in the usual way

$$\chi_{\text{SN}}^2 = (\vec{\mu}_{\text{obs}} - \vec{\mu}_{\text{model}})^T \mathbf{C}_{\text{SN}}^{-1} (\vec{\mu}_{\text{obs}} - \vec{\mu}_{\text{model}}), \quad (18)$$

with \mathbf{C}_{SN} the covariance matrix of $\vec{\mu}_{\text{obs}}$.

3.1.2. GRBs data

The GRBs data we will use is compiled by Amati (2000, 2008, 2009), in which there are 109 long GRBs with measured redshift ($0.1 < z \leq 8.1$) and spectral peak energy. There are 50 GRBs at $z < 1.4$, and 59 GRBs at $z > 1.4$ in this data set, see Ref.(Wei 2010, Table I, II). The well-known Amati correlation (Amati 2002) in GRBs is given by

$$\log_{10} \frac{E_{\text{iso}}}{1 \text{ erg}} = \lambda + b \log_{10} \left(\frac{E_{\text{p,i}}}{300 \text{ KeV}} \right) \quad (19)$$

where E_{iso} is the isotropic-equivalent radiated energy, while $E_{\text{p,i}}$ is the cosmological rest-frame spectral peak energy. Here, λ and b are constants to be determined by observations, see Wei & Cai (2008). The isotropic-equivalent radiated energy E_{iso} is related to the bolometric fluence S_{bolo} of gamma rays in the GRB at redshift z :

$$E_{\text{iso}} = 4\pi D_L^2 S_{\text{bolo}} (1+z)^{-1}. \quad (20)$$

Then, from the GRBs data one can obtain the distance modulus as:

$$\mu_g = \frac{5}{2} \log_{10} \left[\frac{(1+z)}{4\pi} \left(\frac{E_{\text{p,i}}}{300 \text{ KeV}} \right)^b \frac{S_{\text{bolo}}^{-1}}{100 \text{ pc}^2} \right] + \frac{5\lambda}{2}, \quad (21)$$

with uncertainties

$$\sigma_{\mu_g}^2 = \left(\frac{5}{2 \ln 10} \right)^2 \left[b^2 \left(\frac{\sigma_{E_{\text{p,i}}}}{E_{\text{p,i}}} \right)^2 + \left(\frac{\sigma_{S_{\text{bolo}}}}{S_{\text{bolo}}} \right)^2 + \sigma_{\text{sys}}^2 \right]. \quad (22)$$

The χ^2 statistic is then calculated by

$$\chi_g^2 = \sum_{i=1}^N \frac{(\mu_g - \mu_{\text{model}})^2}{\sigma_{\mu_g}^2}, \quad (23)$$

with N data points. Here σ_{sys} in Eq.(22) denotes the systematic error, which accounts the extra scatter of the luminosity relation.

TABLE 1
SYSTEM ERROR EVALUATIONS WITH ITS
STANDARD DEVIATIONS

# of data	PCA		MDS	
	$\langle \sigma_{\text{sys}} \rangle$	std.	$\langle \sigma_{\text{sys}} \rangle$	std.
20	0.7910	0.1361	0.7637	0.1321
30	0.7445	0.0828	0.7492	0.0849
40	0.7588	0.0690	0.7794	0.0721
50	0.7568	0.0495	0.7561	0.0602
60	0.7627	0.0417	0.7546	0.0445
70	0.7558	0.0351	0.7608	0.0360
80	0.7509	0.0322	0.7581	0.0303
90	0.7498	0.0251	0.7595	0.0242
100	0.7531	0.0161	0.7586	0.0167

In literature, the value of σ_{sys} is often estimated by finding the value such that a χ_g^2 fit to the luminosity calibration curve produces a reduced χ_g^2 of unity, see Ref.(Schaefer 2007). In fact, the systematic error should not depend on the number of data points N . Based on this assumption, we randomly choose a subset of the whole 109 GRBs data set, i.e. $N = 20, 30, \dots, 100$. Then, we find the value of σ_{sys} such that the reduced χ_g^2 is unity. We have performed this procedure for 100 times and averaged the value of σ_{sys} , then presented them in Tab.1. Also, the standard deviations of σ_{sys} is given in the same table. Finally, we obtained averaging systematic error (weighted by the stander deviations) as

$$\sigma_{\text{sys}} = 0.7571, \quad (24)$$

which will be used in the next fitting procedures. Besides, from Tab.1, it is clear that the σ_{sys} depends on the model through the χ_g^2 .

3.1.3. Fitting results

During the fitting procedure, we have set the threshold in Eq.(9) to be $r_p > 99.99\%$ for the *PCA-model*. For the *MDS-model*, we require either $r_p^{(1)} > 99.99\%$ or $r_p^{(2)} > 99.99\%$ satisfied, see Eq.(14). Then, we get two parameters c_1 and c_2 for the *PCA-model*, and one parameter \tilde{c}_1 for the *MDS-model*. For comparison, these two models are fitted to observations by using both the Levenberg-Marquardt (LM) technique and the Markov Chain Monte Carlo (MCMC) method. The current value of Hubble parameter is fixed to be $H_0 = 70.0 \text{ km/s/Mpc}$.

At first, only JLA data is used to fit models. After marginalizing the nuisance parameters of JLA, we obtain $c_1 = 10.37 \pm 1.35$, $c_2 = 0.3617 \pm 0.3157$ with $\chi_{\text{min}}^2/\text{d.o.f.} = 683.001/738$ (LM), while $c_1 = 10.40^{+1.39}_{-1.36}$, $c_2 = 0.3540^{+0.3213}_{-0.3254}$ with $\chi_{\text{min}}^2/\text{d.o.f.} = 683.001/738$ (MCMC) for the *PCA-model*. We obtain $\tilde{c}_1 = 2.342 \pm 0.0163$ with $\chi_{\text{min}}^2/\text{d.o.f.} = 683.942/739$ (LM), while $\tilde{c}_1 = 2.343^{+0.0167}_{-0.0163}$ with $\chi_{\text{min}}^2/\text{d.o.f.} = 683.942/739$ (MCMC) for the *MDS-model*. The contours for parameters c_1, c_2 of the *PCA-model* and their 1-D histograms are plotted in Fig.3.

Next, both JLA and GRBs data are used. The nuisance parameters of JLA is also marginalized since we do not have interest in them. However, the parameters λ and b in the Amati correlation (19) are kept free to see how well the calibration is. We obtain $c_1 = 11.52 \pm 0.85$, $c_2 = 0.0434 \pm 0.1546$, $\lambda = 52.850 \pm 0.041$, $b = 1.600 \pm 0.071$ with $\chi_{\text{min}}^2/\text{d.o.f.} = 787.592/845$ (LM), while $c_1 = 11.58^{+0.83}_{-0.82}$, $c_2 = 0.0329^{+0.1500}_{-0.1519}$, $\lambda = 52.852^{+0.039}_{-0.042}$, $b = 1.606^{+0.070}_{-0.072}$

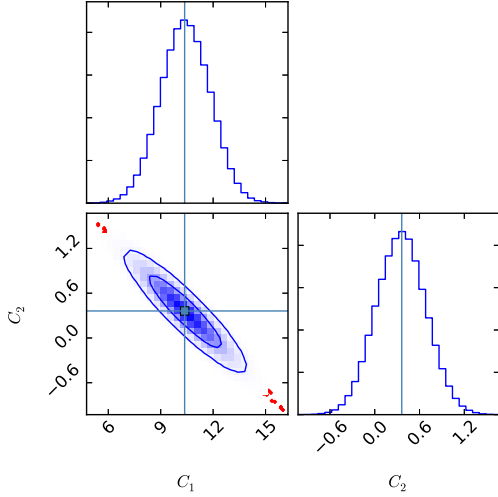


FIG. 3.— JLA data only. The contour from 1σ to 2σ confidence levels and 1-D histograms for parameters c_1, c_2 of the *PCA-model*. The correction between c_1 and c_2 comes from the constraint $E(0) = 1$, see the definition of $E(z)$ below Eq.(1).

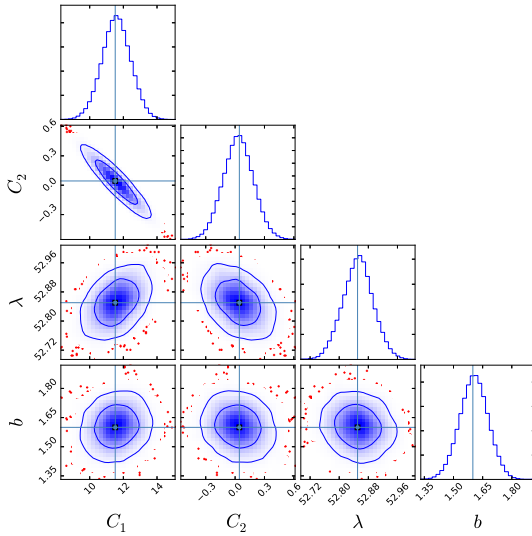


FIG. 4.— JLA + GRBs data. The contours from 1σ to 2σ confidence levels and 1-D histograms for parameters c_1, c_2, λ, b for the *PCA-model*.

with $\chi^2_{\min}/\text{d.o.f.} = 787.601/845$ (MCMC) for the *PCA-model*. We obtain $\tilde{c}_1 = 2.231 \pm 0.016$, $\lambda = 52.841 \pm 0.037$, $b = 1.593 \pm 0.070$ with $\chi^2_{\min}/\text{d.o.f.} = 787.767/846$ (LM), while $\tilde{c}_1 = 2.231^{+0.016}_{-0.015}$, $\lambda = 52.842^{+0.036}_{-0.038}$, $b = 1.590^{+0.074}_{-0.072}$ with $\chi^2_{\min}/\text{d.o.f.} = 787.769/846$ (MCMC) for the *MDS-model*. The contours for parameters c_1, c_2, λ, b and their 1-D histograms are plotted in Fig.4 for the *PCA-model*, while The contours for parameters \tilde{c}_1, λ, b and their 1-D histograms are plotted in Fig.5 for the *MDS-model*.

The calibration of 109 GRBs data is also shown in Fig.6, in which the propagated uncertainties of $\log_{10} E_{\text{iso}}$ and $\log_{10} E_{\text{p},i}$

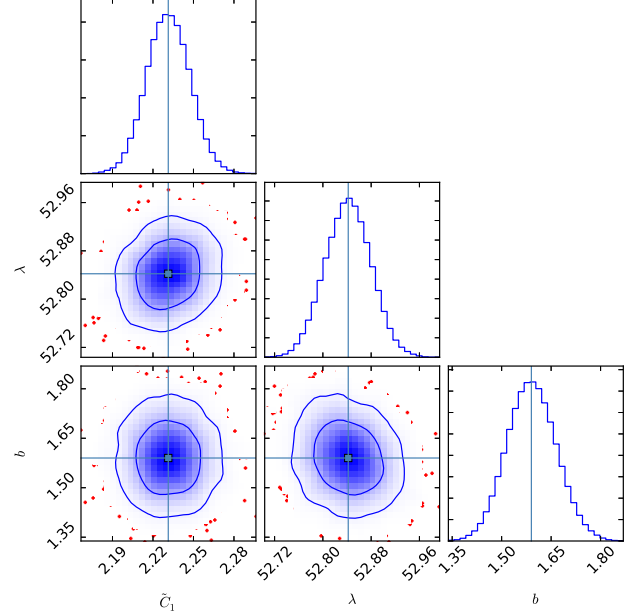


FIG. 5.— JLA + GRBs data. The contours from 1σ to 2σ confidence levels and 1-D histograms for parameters \tilde{c}_1, λ, b for the *MDS-model*.

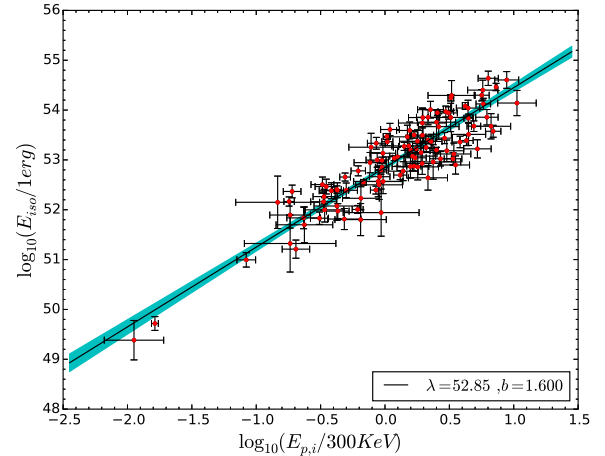


FIG. 6.— Calibration. The red points with error bars corresponds to 109 GRBs data, while the line corresponds the best-fit calibration with 1σ confidence level.

are estimated by

$$\sigma_{\log_{10} E_{\text{iso}}} = \sqrt{\sigma_{\lambda}^2 + \sigma_b^2 \left[\log_{10} \left(\frac{E_{\text{p},i}}{300\text{KeV}} \right) \right]^2 + b^2 \sigma_{\log_{10} E_{\text{p},i}}^2} \quad (25)$$

$$\sigma_{\log_{10} E_{\text{p},i}} = \frac{1}{\ln 10} \frac{\sigma_{E_{\text{p},i}}}{E_{\text{p},i}}. \quad (26)$$

It is clear that the calibration in this work is well consistent with data.

3.2. Reconstruction of History

Now, we are ready to reconstruct the history of Universe. The cosmic expansion rates $E(z) = H(z)/H_0$ with different spatial curvatures are plotted in Fig.7 and Fig.9. The relative

errors of $E(z)$ is estimated by

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{\Omega_k D^2}{1 + \Omega_k D^2}\right)^2 \left(\frac{\sigma_{\Omega_k}^2}{4\Omega_k^2} + \frac{\sigma_D^2}{D^2}\right) + \frac{\sigma_{D'}^2}{D'^2}}, \quad (27)$$

where $\sigma_D = \sqrt{\mathbf{w}_i^2 \sigma_{c_i}^2}$, $\sigma_{D'} = \sqrt{\mathbf{w}_i'^2 \sigma_{c_i}^2}$ for the PCA-model, and $\sigma_D = \sqrt{\tilde{\mathbf{w}}_i^2 \sigma_{\tilde{c}_i}^2}$, $\sigma_{D'} = \sqrt{\tilde{\mathbf{w}}_i'^2 \sigma_{\tilde{c}_i}^2}$ for the MDS-model.

From Fig.7 and Fig.9, one can see that the relative error of $E(z)$ in the MDS-model is about ten times less than that in the PCA-model. And that is as it should be, because there is one parameter \tilde{c}_1 in the MDS-model, while there are two parameters c_1 and c_2 in the PCA-model. In both models, the relative error of $E(z)$ is small at low redshifts, say $0.5 < z < 1.0$, since most of the data points belong to this range of redshifts.

It is interesting to see that in the MDS-model, the relative error of $E(z)$ is a constant for a spatial-flat Universe ($\Omega_k = 0$). This could be seen from Eq.(27): $\sigma_E/E = \sigma_{D'}/D' = \sigma_{\tilde{c}_1}/\tilde{c}_1$, since there is only one parameter \tilde{c}_1 in the MDS-model. Taking the best fitting value for \tilde{c}_1 and its uncertainty $\sigma_{\tilde{c}_1}$, we obtain $\sigma_E/E \approx 0.7\%$.

In Fig.7 and Fig.9, the spatial curvature Ω_k is chosen to show the differences of the cosmic expansion rate under different space geometries. In fact, Li et al. (2014) have already taken a model-independent approach to determine the spatial curvature by using the recent baryon acoustic oscillation (BAO) measurements. According to their conclusions, the errors of Ω_k decrease with increasing redshift, and the best constraint is $\Omega_k = -0.05 \pm 0.06$ (at $z = 2.36$). However the errors of curvature at low redshifts are nearly of order unit, see Li et al. (2014, Fig.2). Considering the future BAO measurements, at least one order of magnitude improvement of Ω_k could be expected at both low and high redshifts (Li et al. 2014).

The ratio of the cosmic expansion rate that predicted from the Λ CDM, the wCDM and the CPL model with their best fitting parameters in Ref. (Benitez-Herrera et al. 2013) to that reconstructed from the PCA-model and the MDS-model, i.e. $H(z)/H_{PCA}(z)$ and $H(z)/H_{MDS}(z)$ are plotted in Fig.8 and Fig.10 respectively.

From Fig.8, one can see that in the Λ CDM model the expansion rate is always smaller than that in the PCA-model, while in the wCDM model $H(z)$ is firstly smaller than $H_{PCA}(z)$ at low redshifts, then becomes larger than $H_{PCA}(z)$ at medium redshifts, and finally gets smaller than $H_{PCA}(z)$ again at high redshifts. In the CPL model, the behavior of $H(z)$ is almost like that in the wCDM model except that $H(z)$ is firstly larger than $H_{PCA}(z)$. From Fig.10, one can see that the behavior of $H(z)$ in these three physical models are almost the same as each other except a small difference at very low redshifts, and they are larger than $H_{MDS}(z)$ at a large range of the redshifts.

Due to the precision limit, we can not find out these differences discussed above from the present observations, since these differences are really quite small.

4. DISCUSSIONS AND CONCLUSIONS

Cosmological variables such as the luminosity distance can be decomposed into some suitable basis. In this paper, we have proposed two methods: PCA and MDS to optimize this basis. The projected feature spaces that describe the luminosity distance could then retain most of the origin information in a low-dimensional space. We call them the PCA-

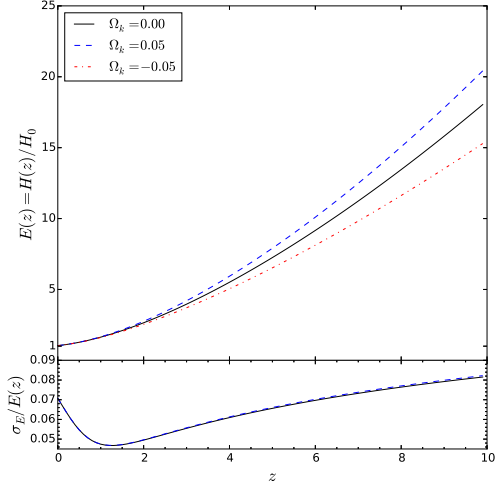


FIG. 7.— Evolution history of the Universe. Top: the reconstructed cosmic expansion rate from the PCA-model. Bottom: relative errors of $E(z)$.

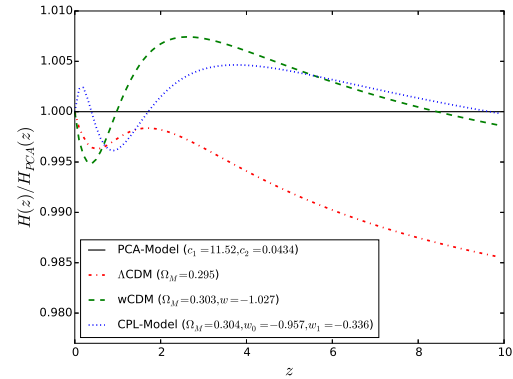


FIG. 8.— Comparison of the cosmic expansion rate predicted from the Λ CDM, the wCDM and the CPL model with their best fitting parameters in Ref. (Benitez-Herrera et al. 2013) and that reconstructed from the PCA-model, namely, $H(z)/H_{PCA}(z)$.

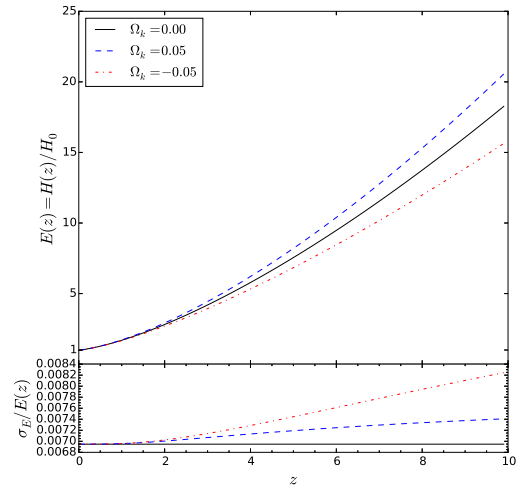


FIG. 9.— Evolution history of the Universe. Top: the reconstructed cosmic expansion rate from the MDS-model. Bottom: relative errors of $E(z)$.

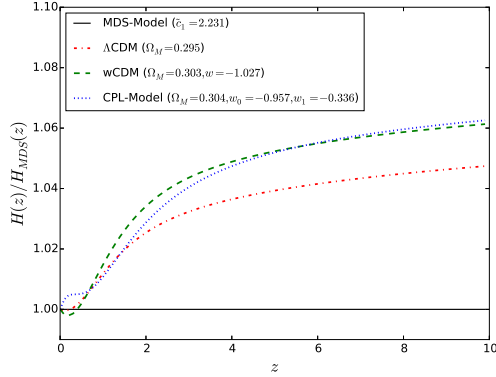


FIG. 10.— Comparison of the cosmic expansion rate predicted from the Λ CDM, the wCDM and the CPL model with their best fitting parameters in Ref. (Benitez-Herrera et al. 2013) and that reconstructed from the MDS-model, namely, $H(z)/H_{MDS}(z)$.

model and the *MDS-model* respectively. It should be noticed that the procedures used above do not depend on any specific cosmological models. After that, observational data including the "joint light-curve analysis" (JLA) data set that consists of 740 Type Ia supernovae (SNIa) as well as 109 long gamma-ray bursts with the well-known Amati relation are

used to constrain the parameters of these two models by using the Levenberg-Marquardt technique and the Markov Chain Monte Carlo method. Finally, we obtain the evolution history of the Universe including both the cosmic expansion rate and its relative errors and we also compare the results with that predicted from the Λ CDM, the wCDM and the CPL model with their best fitting parameters.

We notice that whether the *PCA-model* or the *MDS-model* could be used to perform the calibration to GRBs data without any prior assumptions of a specific cosmological model. We also estimate the system errors of GRBs data. We can say with confidence that the error bars will become smaller when more accurate GRBs data would be obtained in the future.

CJF would like to thank Puxun Wu for helpful discussions. This work is supported by National Science Foundation of China grant Nos. 11105091 and 11047138, "Chen Guang" project supported by Shanghai Municipal Education Commission and Shanghai Education Development Foundation Grant No. 12CG51, National Education Foundation of China grant No. 2009312711004, Shanghai Natural Science Foundation, China grant No. 10ZR1422000, Key Project of Chinese Ministry of Education grant, No. 211059, and Shanghai Special Education Foundation, No. ssd10004, and the Program of Shanghai Normal University.

REFERENCES

- Amati, L. [arXiv:1002.2232]
 Amati, L. et al., 2002, *Astron. Astrophys.*, 390, 81
 Amati, L. et al., 2008, *Mon. Not. Roy. Astron. Soc.*, 391, 577
 Amati, L., Frontera, F. and Guidorzi, C. [arXiv:0907.0384]
 Benitez-Herrera, S., Röpke, F., Hillebrandt, W., et al. 2012, *MNRAS*, 419, 513
 Benitez-Herrera, S., Ishida, E. E. O., Maturi, M., et al. 2013, *MNRAS*, 436, 854
 Betoule, M., et al., 2014, [SDSS Collaboration], *A&A*, 568, A22
 Borg, I., Groenen-Partick, J. F., Mair, P., 2013 *Applied Multidimensional Scaling* (New York: Springer)
 Chevallier, M., & Polarski, D., 2001, *Int. J. Mod. Phys. D*, 10, 213
 Crittenden, R. G., Pogosian, L., & Zhao, G. B., 2009, *J. Cosmology Astropart. Phys.*, 0912, 025
 Dvali, G. R., Gabadadze, G. & Porrati, M., 2000, *Phys. Lett. B*, 485, 208
 Deffayet, C., 2001, *Phys. Lett. B*, 502, 199
 Deffayet, C., Dvali, G. R., & Gabadadze, G., 2002, *Phys. Rev. D*, 65, 044023
 Feng, C. J. & Li, X. Z., 2009, *Phys. Lett. B*, 680, 355
 Feng, C. J., Shen, X.-Y., Li, P., & Li, X. Z., 2012, *J. Cosmology Astropart. Phys.*, 1209, 023
 Firmani, C., Ghisellini, G., Ghirlanda, G., & Avila-Reese, V., 2005, *Mon. Not. Roy. Astron. Soc.* 360, 1
 Gao, C., Chen, X., & Shen, Y. G., 2009, *Phys. Rev. D* 79, 043511
 Ghirlanda, G., Ghisellini, G., Lazzati, D., & Firmani, C., 2004, *AJ*, 613, 13
 Huterer, D., & Starkman, G. 2003, *Phys. Rev. Lett.*, 90, 031301
 Jolliffe, I. T. 2002, *Principal Component Analysis* (2nd ed.; New York: Springer-Verlag)
 Kodama, Y. et al., 2008, *MNRAS*, 391, L1
 Li, H., et al. 2008, *AJ*, 680, 92
 Li, M., 2004, *Phys. Lett. B* 603, 1
 Li, M., Xiao-Dong, L., Shuang, W., & Yi, W., 2011, *Comm. Theor. Phys.*, 56(3), 525
 Li, Y. L., Li, S. Y., Zhang, T. J., Li, T. P., 2014, *AJ*, 789, 15
 Liang, N., Xiao, W. K., Liu, Y., & Zhang, S. N., 2008, *AJ*, 685, 354
 Linder, E. V., 2003, *Phys. Rev. Lett.* 90, 091301
 Maturi, M., & Mignone, C., 2009, *A&A*, 508, 45
 Mignone, C. and Bartelmann, M. 2008, *A&A*, 481, 295
 Perlmutter, S., et al. 1999, *ApJ*, 517, 565
 Riess, A. G., et al. 1998, *AJ*, 116, 1009
 Schaefer, B. E., 2007, *AJ*, 660, 16
 Shafer, D. L., 2015, *Phys. Rev. D* 91, no. 10, 103516
 Simpson, F., & Bridle, S. 2006, *Phys. Rev. D*, 73, 083001
 Wei, H., & Cai, R. G., 2008, *Phys. Lett. B* 660, 113
 Wei, H., 2010, *J. Cosmology Astropart. Phys.*, 1008, 020

TABLE 2
 THE EFFICIENCY OF PCA AND MDS FOR DIFFERENT TSGs

TSGs	PCA		MDS				Parameters
	$\langle r_1 \rangle$	$\langle r_2 \rangle$	$\langle r_1^{(1)} \rangle$	$\langle r_2^{(1)} \rangle$	$\langle r_1^{(2)} \rangle$	$\langle r_2^{(2)} \rangle$	
Λ CDM	99.834%	99.996%	99.692%	99.998%	99.998%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, -0.1 < \Omega_K < 0.1$
wCDM	99.859%	99.995%	99.413%	99.992%	99.995%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, -0.1 < \Omega_K < 0.1, -1.5 < w < -0.5$
CPL ^a	99.906%	99.998%	99.701%	99.997%	99.998%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, -1.5 < w_0 < -0.5, -0.5 < w_1 < 0.5$
FSLI-I ^b	99.895%	99.997%	99.767%	99.997%	99.998%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, -1.5 < w_0 < -0.5, -0.5 < w_1 < 0.5$
FSLI-II ^b	99.905%	99.998%	99.795%	99.998%	99.998%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, -1.5 < w_0 < -0.5, -0.5 < w_1 < 0.5$
HDE ^c	99.922%	99.999%	99.460%	99.996%	99.997%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, 0.1 < C < 1.5$
DGP ^d	99.804%	99.999%	99.260%	99.998%	99.992%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9$
NADE ^e	99.897%	99.999%	99.706%	99.998%	99.998%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, 1.5 < n < 3.5$
RDE ^f	99.947%	99.998%	99.848%	99.995%	99.998%	99.999%	$0.1 < \Omega_M < 0.9, 0.1 < \Omega_V < 0.9, 0.1 < \alpha < 1.0$
Mixture	99.780%	99.995%	99.250%	99.998%	99.995%	99.999%	Take the same ranges as above.

^a Chevallier & Polarski (2001); Linder (2003)

^b Feng et al. (2012)

^c Li (2004)

^d Dvali et al. (2000); Deffayet (2001, 2002)

^e Wei & Cai (2008)

^f Gao et al. (2009); Feng & Li (2009)